

SEQUENCE LISTING

(1) GENERAL INFORMATION:

- (i) APPLICANT: Gehrmann, Mathias
Seemann, Gerhard
Bosslet, Klaus
Czech, Joerg
- (ii) TITLE OF INVENTION: Fusion Protein for Prodrug Activity
- (iii) NUMBER OF SEQUENCES: 18
- (iv) CORRESPONDENCE ADDRESS:
 - (A) ADDRESSEE: Finnegan, Henderson, Farabow, Garrett & Dunner
 - (B) STREET: 1300 I Street, N.W.
 - (C) CITY: Washington
 - (D) STATE: D.C.
 - (E) COUNTRY: USA
 - (F) ZIP: 20005-3315
- (v) COMPUTER READABLE FORM:
 - (A) MEDIUM TYPE: Floppy disk
 - (B) COMPUTER: IBM PC compatible
 - (C) OPERATING SYSTEM: PC-DOS/MS-DOS
 - (D) SOFTWARE: PatentIn Release #1.0, Version #1.25
- (vi) CURRENT APPLICATION DATA:
 - (A) APPLICATION NUMBER: US 08/129,379
 - (B) FILING DATE: 30-SEP-1993
 - (C) CLASSIFICATION:
- (vii) PRIOR APPLICATION DATA:
 - (A) APPLICATION NUMBER: DE P 42 33 152.8
 - (B) FILING DATE: 02-OCT-1992
- (viii) ATTORNEY/AGENT INFORMATION:
 - (A) NAME: Barker, M. P.
 - (B) REGISTRATION NUMBER: 32,013
 - (C) REFERENCE/DOCKET NUMBER: 02481.1337-00000
- (ix) TELECOMMUNICATION INFORMATION:
 - (A) TELEPHONE: 202-408-4000
 - (B) TELEFAX: 202-408-4400

(2) INFORMATION FOR SEQ ID NO:1:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 3314 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:

(A) NAME/KEY: sig_peptide

(B) LOCATION: 145..283

(ix) FEATURE:

(A) NAME/KEY: mat_peptide

(B) LOCATION: join(284..1003, 1069..1119, 1263..3161)

(ix) FEATURE:

(A) NAME/KEY: CDS

(B) LOCATION: join(145..189, 272..1003, 1069..1119, 1263..3161)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

CCAAGCTTAT GAATATGCAA ATCCTGCTCA TGAATATGCA AATCCTCTGA ATCTACATGG	60
TAAATATAGG TTTGTCTATA CCACAAACAG AAAAACATGA GATCACAGTT CTCTCTACAG	120
TTACTGAGCA CACAGGACCT CACC ATG GGA TGG AGC TGT ATC ATC CTC TTC	171
Met Gly Trp Ser Cys Ile Ile Leu Phe	
-19 -15	
TTG GTA GCA ACA GCT ACA GGTAAGGGGC TCACAGTAGC AGGCTTGAGG	219
Leu Val Ala Thr Ala Thr	
-10 -5	
TCTGGACATA TATATGGGTG ACAATGACAT CCACTTTGCC TTTCTCTCCA CA GGT	274
Gly	
-4	
GTC CAC TCC CAG GTC CAA CTG CAG GAG AGC GGT CCA GGT CTT GTG AGA	322
Val His Ser Gln Val Gln Leu Gln Glu Ser Gly Pro Gly Leu Val Arg	
1 5 10	
CCT AGC CAG ACC CTG AGC CTG ACC TGC ACC GTG TCT GGC TTC ACC ATC	370
Pro Ser Gln Thr Leu Ser Leu Thr Cys Thr Val Ser Gly Phe Thr Ile	
15 20 25	
AGC AGT GGT TAT AGC TGG CAC TGG GTG AGA CAG CCA CCT GGA CGA GGT	418
Ser Ser Gly Tyr Ser Trp His Trp Val Arg Gln Pro Pro Gly Arg Gly	
30 35 40 45	
CTT GAG TGG ATT GGA TAC ATA CAG TAC AGT GGT ATC ACT AAC TAC AAC	466
Leu Glu Trp Ile Gly Tyr Ile Gln Tyr Ser Gly Ile Thr Asn Tyr Asn	
50 55 60	

CCC Pro	TCT Ser	CTC Leu	AAA Lys 65	AGT Ser	AGA Arg	GTG Val	ACA Thr	ATG Met 70	CTG Leu	GTA Val	GAC Asp	ACC Thr	AGC Ser 75	AAG Lys	AAC Asn	514
CAG Gln	TTC Phe	AGC Ser 80	CTG Leu	AGA Arg	CTC Leu	AGC Ser	AGC Ser 85	GTG Val	ACA Thr	GCC Ala	GCC Ala	GAC Asp 90	ACC Thr	GCG Ala	GTC Val	562
TAT Tyr	TAT Tyr 95	TGT Cys	GCA Ala	AGA Arg	GAA Glu	GAC Asp 100	TAT Tyr	GAT Asp	TAC Tyr	CAC His	TGG Trp 105	TAC Tyr	TTC Phe	GAT Asp	GTC Val	610
TGG Trp 110	GGC Gly	CAA Gln	GGG Gly	ACC Thr	ACG Thr 115	GTC Val	ACC Thr	GTC Val	TCC Ser	TCA Ser 120	GGA Gly	GGC Gly	GGT Gly	GGA Gly	TCG Ser 125	658
GGC Gly	GGT Gly	GGT Gly	GGG Gly	TCG Ser 130	GGT Gly	GGC Gly	GGC Gly	GGA Gly	TCT Ser 135	GAC Asp	ATC Ile	CAG Gln	CTG Leu	ACC Thr 140	CAG Gln	706
AGC Ser	CCA Pro	AGC Ser	AGC Ser 145	CTG Leu	AGC Ser	GCC Ala	AGC Ser	GTG Val 150	GGT Gly	GAC Asp	AGA Arg	GTG Val 155	ACC Thr	ATC Ile	ACC Thr	754
TGT Cys	AGT Ser	ACC Thr 160	AGC Ser	TCG Ser	AGT Ser	GTA Val	AGT Ser 165	TAC Tyr	ATG Met	CAC His	TGG Trp	TAC Tyr 170	CAG Gln	CAG Gln	AAG Lys	802
CCA Pro	GGT Gly 175	AAG Lys	GCT Ala	CCA Pro	AAG Lys	CTG Leu 180	CTG Leu	ATC Ile	TAC Tyr	AGC Ser	ACA Thr 185	TCC Ser	AAC Asn	CTG Leu	GCT Ala	850
TCT Ser 190	GGT Gly	GTG Val	CCA Pro	AGC Ser	AGA Arg 195	TTC Phe	AGC Ser	GGT Gly	AGC Ser	GGT Gly 200	AGC Ser	GGT Gly	ACC Thr	GAC Asp	TTC Phe 205	898
ACC Thr	TTC Phe	ACC Thr	ATC Ile	AGC Ser 210	AGC Ser	CTC Leu	CAG Gln	CCA Pro	GAG Glu 215	GAC Asp	ATC Ile	GCC Ala	ACC Thr	TAC Tyr 220	TAC Tyr	946
TGC Cys	CAT His	CAG Gln	TGG Trp 225	AGT Ser	AGT Ser	TAT Tyr	CCC Pro	ACG Thr 230	TTC Phe	GGC Gly	CAA Gln	GGG Gly 235	ACC Thr	AAG Lys	CTG Leu	994
GAG Glu	ATC Ile	AAA Lys 240	GGTGAGTAGA				ATTTAAACTT			TGCTTCCTCA		GTTGGATCTG				1043
AGTAACTCCC			AATCTTCTCT			CTGCA	GAG Glu	CTC Leu	AAA Lys	ACC Thr	CCA Pro 245	CTT Leu	GGT Gly	GAC Asp	ACA Thr	1095

ACT CAC ACA TGC CCA CGG TGC CCA GGTAAGCCAG CCCAGGACTC GCCCTCCAGC	1149
Thr His Thr Cys Pro Arg Cys Pro	
250 255	
TCAAGGCGGG ACAAGAGCCC TAGAGTGGCC TGAGTCCAGG GACAGGCCCC AGCAGGGTGC	1209
TGACGCATCC ACCTCCATCC CAGATCCCCG TAACTCCCAA TCTTCTCTCT GCA GCG	1265
Ala	
GCG GCG GCG GTG CAG GGC GGG ATG CTG TAC CCC CAG GAG AGC CCG TCG	1313
Ala Ala Ala Val Gln Gly Gly Met Leu Tyr Pro Gln Glu Ser Pro Ser	
260 265 270	
CGG GAG TGC AAG GAG CTG GAC GGC CTC TGG AGC TTC CGC GCC GAC TTC	1361
Arg Glu Cys Lys Glu Leu Asp Gly Leu Trp Ser Phe Arg Ala Asp Phe	
275 280 285 290	
TCT GAC AAC CGA CGC CGG GGC TTC GAG GAG CAG TGG TAC CGG CGG CCG	1409
Ser Asp Asn Arg Arg Arg Gly Phe Glu Glu Gln Trp Tyr Arg Arg Pro	
295 300 305	
CTG TGG GAG TCA GGC CCC ACC GTG GAC ATG CCA GTT CCC TCC AGC TTC	1457
Leu Trp Glu Ser Gly Pro Thr Val Asp Met Pro Val Pro Ser Ser Phe	
310 315 320	
AAT GAC ATC AGC CAG GAC TGG CGT CTG CGG CAT TTT GTC GGC TGG GTG	1505
Asn Asp Ile Ser Gln Asp Trp Arg Leu Arg His Phe Val Gly Trp Val	
325 330 335	
TGG TAC GAA CGG GAG GTG ATC CTG CCG GAG CGA TGG ACC CAG GAC CTG	1553
Trp Tyr Glu Arg Glu Val Ile Leu Pro Glu Arg Trp Thr Gln Asp Leu	
340 345 350	
CGC ACA AGA GTG GTG CTG AGG ATT GGC AGT GCC CAT TCC TAT GCC ATC	1601
Arg Thr Arg Val Val Leu Arg Ile Gly Ser Ala His Ser Tyr Ala Ile	
355 360 365 370	
GTG TGG GTG AAT GGG GTC GAC ACG CTA GAG CAT GAG GGG GGC TAC CTC	1649
Val Trp Val Asn Gly Val Asp Thr Leu Glu His Glu Gly Gly Tyr Leu	
375 380 385	
CCC TTC GAG GCC GAC ATC AGC AAC CTG GTC CAG GTG GGG CCC CTG CCC	1697
Pro Phe Glu Ala Asp Ile Ser Asn Leu Val Gln Val Gly Pro Leu Pro	
390 395 400	
TCC CGG CTC CGA ATC ACT ATC GCC ATC AAC AAC ACA CTC ACC CCC ACC	1745
Ser Arg Leu Arg Ile Thr Ile Ala Ile Asn Asn Thr Leu Thr Pro Thr	
405 410 415	
ACC CTG CCA CCA GGG ACC ATC CAA TAC CTG ACT GAC ACC TCC AAG TAT	1793
Thr Leu Pro Pro Gly Thr Ile Gln Tyr Leu Thr Asp Thr Ser Lys Tyr	
420 425 430	

CCC Pro 435	AAG Lys	GGT Gly	TAC Tyr	TTT Phe	GTC Val 440	CAG Gln	AAC Asn	ACA Thr	TAT Tyr	TTT Phe 445	GAC Asp	TTT Phe	TTC Phe	AAC Asn	TAC Tyr 450	1841
GCT Ala	GGA Gly	CTG Leu	CAG Gln	CGG Arg 455	TCT Ser	GTA Val	CTT Leu	CTG Leu	TAC Tyr 460	ACG Thr	ACA Thr	CCC Pro	ACC Thr	ACC Thr 465	TAC Tyr	1889
ATC Ile	GAT Asp	GAC Asp	ATC Ile 470	ACC Thr	GTC Val	ACC Thr	ACC Thr	AGC Ser 475	GTG Val	GAG Glu	CAA Gln	GAC Asp	AGT Ser 480	GGG Gly	CTG Leu	1937
GTG Val	AAT Asn 485	TAC Tyr	CAG Gln	ATC Ile	TCT Ser	GTC Val	AAG Lys 490	GGC Gly	AGT Ser	AAC Asn	CTG Leu	TTC Phe 495	AAG Lys	TTG Leu	GAA Glu	1985
GTG Val 500	CGT Arg	CTT Leu	TTG Leu	GAT Asp	GCA Ala 505	GAA Glu 505	AAC Asn	AAA Lys	GTC Val	GTG Val	GCG Ala 510	AAT Asn	GGG Gly	ACT Thr	GGG Gly	2033
ACC Thr 515	CAG Gln	GGC Gly	CAA Gln	CTT Leu	AAG Lys 520	GTG Val	CCA Pro	GGT Gly	GTC Val	AGC Ser 525	CTC Leu	TGG Trp	TGG Trp	CCG Pro	TAC Tyr 530	2081
CTG Leu	ATG Met	CAC His	GAA Glu	CGC Arg 535	CCT Pro	GCC Ala	TAT Tyr	CTG Leu	TAT Tyr 540	TCA Ser	TTG Leu	GAG Glu	GTG Val	CAG Gln 545	CTG Leu	2129
ACT Thr	GCA Ala	CAG Gln	ACG Thr 550	TCA Ser	CTG Leu	GGG Gly	CCT Pro	GTG Val 555	TCT Ser	GAC Asp	TTC Phe	TAC Tyr	ACA Thr 560	CTC Leu	CCT Pro	2177
GTG Val	GGG Gly 565	ATC Ile	CGC Arg	ACT Thr	GTG Val	GCT Ala	GTC Val 570	ACC Thr	AAG Lys	AGC Ser	CAG Gln	TTC Phe 575	CTC Leu	ATC Ile	AAT Asn	2225
GGG Gly 580	AAA Lys	CCT Pro	TTC Phe	TAT Tyr	TTC Phe	CAC His 585	GGT Gly	GTC Val	AAC Asn	AAG Lys	CAT His 590	GAG Glu	GAT Asp	GCG Ala	GAC Asp	2273
ATC Ile 595	CGA Arg	GGG Gly	AAG Lys	GGC Gly	TTC Phe 600	GAC Asp	TGG Trp	CCG Pro	CTG Leu	CTG Leu 605	GTG Val	AAG Lys	GAC Asp	TTC Phe	AAC Asn 610	2321
CTG Leu	CTT Leu	CGC Arg	TGG Trp	CTT Leu 615	GGT Gly	GCC Ala	AAC Asn	GCT Ala	TTC Phe 620	CGT Arg	ACC Thr	AGC Ser	CAC His	TAC Tyr 625	CCC Pro	2369
TAT Tyr	GCA Ala	GAG Glu	GAA Glu 630	GTG Val	ATG Met	CAG Gln	ATG Met	TGT Cys 635	GAC Asp	CGC Arg	TAT Tyr	GGG Gly	ATT Ile 640	GTG Val	GTC Val	2417

ATC Ile	GAT Asp	GAG Glu 645	TGT Cys	CCC Pro	GGC Gly	GTG Val 650	GGC Gly 650	CTG Leu	GCG Ala	CTG Leu	CCG Pro	CAG Gln 655	TTC Phe	TTC Phe	AAC Asn	2465
AAC Asn 660	GTT Val	TCT Ser	CTG Leu	CAT His	CAC His 665	CAC His 665	ATG Met	CAG Gln	GTG Val	ATG Met	GAA Glu 670	GAA Glu	GTG Val	GTG Val	CGT Arg	2513
AGG Arg 675	GAC Asp	AAG Lys	AAC Asn	CAC His 680	CCC Pro 680	GCG Ala	GTC Val	GTG Val	ATG Met	TGG Trp 685	TCT Ser	GTG Val	GCC Ala	AAC Asn	GAG Glu 690	2561
CCT Pro	GCG Ala	TCC Ser	CAC His 695	CTA Leu 695	GAA Glu	TCT Ser	GCT Ala	GGC Gly 700	TAC Tyr 700	TAC Tyr	TTG Leu	AAG Lys	ATG Met	GTG Val 705	ATC Ile	2609
GCT Ala	CAC His	ACC Thr 710	AAA Lys 710	TCC Ser	TTG Leu	GAC Asp	CCC Pro 715	TCC Ser 715	CGG Arg	CCT Pro	GTG Val	ACC Thr 720	TTT Phe 720	GTG Val	AGC Ser	2657
AAC Asn 725	TCT Ser	AAC Asn 725	TAT Tyr	GCA Ala	GCA Ala	GAC Asp 730	AAG Lys 730	GGG Gly	GCT Ala	CCG Pro	TAT Tyr	GTG Val 735	GAT Asp	GTG Val	ATC Ile	2705
TGT Cys 740	TTG Leu	AAC Asn	AGC Ser	TAC Tyr	TAC Tyr	TCT Ser 745	TGG Trp	TAT Tyr	CAC His	GAC Asp 750	TAC Tyr	GGG Gly	CAC His	CTG Leu	GAG Glu	2753
TTG Leu 755	ATT Ile	CAG Gln	CTG Leu	CAG Gln 760	CTG Leu 760	GCC Ala	ACC Thr	CAG Gln	TTT Phe	GAG Glu 765	AAC Asn	TGG Trp	TAT Tyr	AAG Lys 770	AAG Lys 770	2801
TAT Tyr	CAG Gln	AAG Lys	CCC Pro	ATT Ile 775	ATT Ile 775	CAG Gln	AGC Ser	GAG Glu	TAT Tyr 780	GGA Gly	GCA Ala	GAA Glu	ACG Thr	ATT Ile 785	GCA Ala 785	2849
GGG Gly	TTT Phe	CAC His 790	CAG Gln	GAT Asp	CCA Pro	CCT Pro	CTG Leu	ATG Met 795	TTC Phe	ACT Thr	GAA Glu	GAG Glu	TAC Tyr 800	CAG Gln	AAA Lys	2897
AGT Ser	CTG Leu	CTA Leu 805	GAG Glu	CAG Gln	TAC Tyr	CAT His 810	CTG Leu 810	GGT Gly	CTG Leu	GAT Asp	CAA Gln	AAA Lys 815	CGC Arg	AGA Arg	AAA Lys	2945
TAT Tyr 820	GTG Val	GTT Val	GGA Gly	GAG Glu	CTC Leu	ATT Ile 825	TGG Trp	AAT Asn	TTT Phe	GCC Ala 830	GAT Asp	TTC Phe	ATG Met	ACT Thr	GAA Glu	2993
CAG Gln 835	TCA Ser	CCG Pro	ACG Thr	AGA Arg	GTG Val 840	CTG Leu	GGG Gly	ATT Asn	AAA Lys	AAG Lys 845	GGG Gly	ATC Ile	TTC Phe	ACT Thr	CGG Arg 850	3041

CAG AGA CAA CCA AAA AGT GCA GCG TTC CTT TTG CGA GAG AGA TAC TGG	3089
Gln Arg Gln Pro Lys Ser Ala Ala Phe Leu Leu Arg Glu Arg Tyr Trp	
855 860 865	
AAG ATT GCC AAT GAA ACC AGG TAT CCC CAC TCA GTA GCC AAG TCA CAA	3137
Lys Ile Ala Asn Glu Thr Arg Tyr Pro His Ser Val Ala Lys Ser Gln	
870 875 880	
TGT TTG GAA AAC AGC CCG TTT ACT TGAGCAAGAC TGATAACCACC TGC GTGTCCC	3191
Cys Leu Glu Asn Ser Pro Phe Thr	
885 890	
TTCCTCCCCG AGTCAGGGCG ACTTCCACAG CAGCAGAACA AGTGCCTCCT GGACTGTTC	3251
CGGCAGACCA GAACGTTTCT GGCCTGGGT TTGTGGTCAT CTATTCTAGC AGGGAACACT	3311
AAA	3314

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 909 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Met Gly Trp Ser Cys Ile Ile Leu Phe Leu Val Ala Thr Ala Thr Gly	
-19 -15 -10 -5	
Val His Ser Gln Val Gln Leu Gln Glu Ser Gly Pro Gly Leu Val Arg	
1 5 10	
Pro Ser Gln Thr Leu Ser Leu Thr Cys Thr Val Ser Gly Phe Thr Ile	
15 20 25	
Ser Ser Gly Tyr Ser Trp His Trp Val Arg Gln Pro Pro Gly Arg Gly	
30 35 40 45	
Leu Glu Trp Ile Gly Tyr Ile Gln Tyr Ser Gly Ile Thr Asn Tyr Asn	
50 55 60	
Pro Ser Leu Lys Ser Arg Val Thr Met Leu Val Asp Thr Ser Lys Asn	
65 70 75	
Gln Phe Ser Leu Arg Leu Ser Ser Val Thr Ala Ala Asp Thr Ala Val	
80 85 90	
Tyr Tyr Cys Ala Arg Glu Asp Tyr Asp Tyr His Trp Tyr Phe Asp Val	
95 100 105	

Trp 110	Gly	Gln	Gly	Thr	Thr	Val	Thr	Val	Ser	Ser	Gly	Gly	Gly	Gly	Ser 125
Gly	Gly	Gly	Gly	Ser 130	Gly	Gly	Gly	Gly	Ser 135	Asp	Ile	Gln	Leu	Thr	Gln 140
Ser	Pro	Ser	Ser 145	Leu	Ser	Ala	Ser	Val 150	Gly	Asp	Arg	Val	Thr 155	Ile	Thr
Cys	Ser	Thr 160	Ser	Ser	Ser	Val	Ser 165	Tyr	Met	His	Trp	Tyr 170	Gln	Gln	Lys
Pro	Gly 175	Lys	Ala	Pro	Lys	Leu 180	Leu	Ile	Tyr	Ser	Thr 185	Ser	Asn	Leu	Ala
Ser 190	Gly	Val	Pro	Ser	Arg 195	Phe	Ser	Gly	Ser 200	Gly	Ser	Gly	Thr	Asp	Phe 205
Thr	Phe	Thr	Ile	Ser 210	Ser	Leu	Gln	Pro	Glu 215	Asp	Ile	Ala	Thr	Tyr 220	Tyr
Cys	His	Gln	Trp 225	Ser	Ser	Tyr	Pro	Thr 230	Phe	Gly	Gln	Gly	Thr 235	Lys	Leu
Glu	Ile 240	Lys	Glu	Leu	Lys	Thr	Pro 245	Leu	Gly	Asp	Thr	Thr 250	His	Thr	Cys
Pro	Arg 255	Cys	Pro	Ala	Ala	Ala	Ala	Val	Gln	Gly	Gly 265	Met	Leu	Tyr	Pro
Gln 270	Glu	Ser	Pro	Ser	Arg 275	Glu	Cys	Lys	Glu	Leu 280	Asp	Gly	Leu	Trp	Ser 285
Phe	Arg	Ala	Asp	Phe 290	Ser	Asp	Asn	Arg	Arg 295	Arg	Gly	Phe	Glu	Glu	Gln 300
Trp	Tyr	Arg	Arg 305	Pro	Leu	Trp	Glu	Ser 310	Gly	Pro	Thr	Val	Asp 315	Met	Pro
Val	Pro 320	Ser	Ser	Phe	Asn	Asp	Ile 325	Ser	Gln	Asp	Trp	Arg 330	Leu	Arg	His
Phe 335	Val	Gly	Trp	Val	Trp	Tyr 340	Glu	Arg	Glu	Val	Ile 345	Leu	Pro	Glu	Arg
Trp 350	Thr	Gln	Asp	Leu	Arg 355	Thr	Arg	Val	Val	Leu 360	Arg	Ile	Gly	Ser	Ala 365
His	Ser	Tyr	Ala	Ile 370	Val	Trp	Val	Asn	Gly 375	Val	Asp	Thr	Leu	Glu	His 380
Glu	Gly	Gly	Tyr 385	Leu	Pro	Phe	Glu	Ala	Asp	Ile	Ser	Asn	Leu 395	Val	Gln

Val	Gly	Pro	Leu	Pro	Ser	Arg	Leu	Arg	Ile	Thr	Ile	Ala	Ile	Asn	Asn	
		400					405					410				
Thr	Leu	Thr	Pro	Thr	Thr	Leu	Pro	Pro	Gly	Thr	Ile	Gln	Tyr	Leu	Thr	
	415					420					425					
Asp	Thr	Ser	Lys	Tyr	Pro	Lys	Gly	Tyr	Phe	Val	Gln	Asn	Thr	Tyr	Phe	
430					435					440					445	
Asp	Phe	Phe	Asn	Tyr	Ala	Gly	Leu	Gln	Arg	Ser	Val	Leu	Leu	Tyr	Thr	
				450					455					460		
Thr	Pro	Thr	Thr	Tyr	Ile	Asp	Asp	Ile	Thr	Val	Thr	Thr	Ser	Val	Glu	
			465					470					475			
Gln	Asp	Ser	Gly	Leu	Val	Asn	Tyr	Gln	Ile	Ser	Val	Lys	Gly	Ser	Asn	
		480					485					490				
Leu	Phe	Lys	Leu	Glu	Val	Arg	Leu	Leu	Asp	Ala	Glu	Asn	Lys	Val	Val	
	495					500					505					
Ala	Asn	Gly	Thr	Gly	Thr	Gln	Gly	Gln	Leu	Lys	Val	Pro	Gly	Val	Ser	
510					515					520					525	
Leu	Trp	Trp	Pro	Tyr	Leu	Met	His	Glu	Arg	Pro	Ala	Tyr	Leu	Tyr	Ser	
				530					535					540		
Leu	Glu	Val	Gln	Leu	Thr	Ala	Gln	Thr	Ser	Leu	Gly	Pro	Val	Ser	Asp	
			545					550					555			
Phe	Tyr	Thr	Leu	Pro	Val	Gly	Ile	Arg	Thr	Val	Ala	Val	Thr	Lys	Ser	
		560					565					570				
Gln	Phe	Leu	Ile	Asn	Gly	Lys	Pro	Phe	Tyr	Phe	His	Gly	Val	Asn	Lys	
	575					580					585					
His	Glu	Asp	Ala	Asp	Ile	Arg	Gly	Lys	Gly	Phe	Asp	Trp	Pro	Leu	Leu	
590					595					600					605	
Val	Lys	Asp	Phe	Asn	Leu	Leu	Arg	Trp	Leu	Gly	Ala	Asn	Ala	Phe	Arg	
				610					615					620		
Thr	Ser	His	Tyr	Pro	Tyr	Ala	Glu	Glu	Val	Met	Gln	Met	Cys	Asp	Arg	
			625					630					635			
Tyr	Gly	Ile	Val	Val	Ile	Asp	Glu	Cys	Pro	Gly	Val	Gly	Leu	Ala	Leu	
		640					645					650				
Pro	Gln	Phe	Phe	Asn	Asn	Val	Ser	Leu	His	His	His	Met	Gln	Val	Met	
	655					660					665					
Glu	Glu	Val	Val	Arg	Arg	Asp	Lys	Asn	His	Pro	Ala	Val	Val	Met	Trp	
670					675					680					685	

Ser	Val	Ala	Asn	Glu	Pro	Ala	Ser	His	Leu	Glu	Ser	Ala	Gly	Tyr	Tyr	690	695	700
Leu	Lys	Met	Val	Ile	Ala	His	Thr	Lys	Ser	Leu	Asp	Pro	Ser	Arg	Pro	705	710	715
Val	Thr	Phe	Val	Ser	Asn	Ser	Asn	Tyr	Ala	Ala	Asp	Lys	Gly	Ala	Pro	720	725	730
Tyr	Val	Asp	Val	Ile	Cys	Leu	Asn	Ser	Tyr	Tyr	Ser	Trp	Tyr	His	Asp	735	740	745
Tyr	Gly	His	Leu	Glu	Leu	Ile	Gln	Leu	Gln	Leu	Ala	Thr	Gln	Phe	Glu	750	755	760
Asn	Trp	Tyr	Lys	Lys	Tyr	Gln	Lys	Pro	Ile	Ile	Gln	Ser	Glu	Tyr	Gly	770	775	780
Ala	Glu	Thr	Ile	Ala	Gly	Phe	His	Gln	Asp	Pro	Pro	Leu	Met	Phe	Thr	785	790	795
Glu	Glu	Tyr	Gln	Lys	Ser	Leu	Leu	Glu	Gln	Tyr	His	Leu	Gly	Leu	Asp	800	805	810
Gln	Lys	Arg	Arg	Lys	Tyr	Val	Val	Gly	Glu	Leu	Ile	Trp	Asn	Phe	Ala	815	820	825
Asp	Phe	Met	Thr	Glu	Gln	Ser	Pro	Thr	Arg	Val	Leu	Gly	Asn	Lys	Lys	830	835	840
Gly	Ile	Phe	Thr	Arg	Gln	Arg	Gln	Pro	Lys	Ser	Ala	Ala	Phe	Leu	Leu	850	855	860
Arg	Glu	Arg	Tyr	Trp	Lys	Ile	Ala	Asn	Glu	Thr	Arg	Tyr	Pro	His	Ser	865	870	875
Val	Ala	Lys	Ser	Gln	Cys	Leu	Glu	Asn	Ser	Pro	Phe	Thr				880	885	890

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 26 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

ACCAGAAGCT TATGAATATG CAAATC

(2) INFORMATION FOR SEQ ID NO:4:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 58 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

GCCACCCGAC CCACCACCGC CCGATCCACC GCCTCCTGAG GAGACGGTGA CCGTGGTC 58

(2) INFORMATION FOR SEQ ID NO:5:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 60 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

GGTGGATCGG GCGGTGGTGG GTCGGGTGGC GGCGGATCTG ACATCCAGCT GACCCAGAGC 60

(2) INFORMATION FOR SEQ ID NO:6:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 50 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

TGCAGGATCC AACTGAGGAA GCAAAGTTTA AATTCTACTC ACCTTTGATC 50

(2) INFORMATION FOR SEQ ID NO:7:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 28 base pairs
 - (B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

TTTTTAAGCT TAGATCTCCA CCTTGGTC

28

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 34 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

AAAAATCTAG AATGCAGGTC CAACTGCAGG AGAG

34

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 48 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

AAAAAAGTGA TCAAAGCGTC TGGCGGGCCA CAGGGCGGGA TCCTGTAC

48

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 27 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

TTTTAAGCTT CAAGTAAACG GGCTGTT

27

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 43 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

TTTTGGTACC TTTGGATAAA AGACAGGTCC AACTGCAGGA GAG

43

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 40 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

AAAACCATGG GAATTCAAGC TTCGAGCTGG TACTACAGGT

40

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 44 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

TTTTAAGCTT CCATGGCGGC CGCTCATTGT TTGCCTCCCT GCTG

44

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 48 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

AAAAAGATCT CCGCGTCTGG CGGGCCACAG TTACGTGTAG AAACCCCA

48

(2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 43 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

TTTGGTACC TTTGGATAAA AGACAGGTCC AACTGCAGGA GAG

43

(2) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 28 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

AAAAAGCTTA GATCTCCAGC TTGGTCCC

28

(2) INFORMATION FOR SEQ ID NO:17:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 55 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

AAAGAATTCT GATCAAATCC TCGAGCTCAG GTTCACAAAA GGTAGAGAAA ACAGT

55

(2) INFORMATION FOR SEQ ID NO:18:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 28 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

TTTAAGCTTA TTTTAATAAA TCCAATGT

28